Chapter 5 - Continuous Random Variables

5.1 - Introduction

Reminder

In Ch 4 we have seen discrete random variables

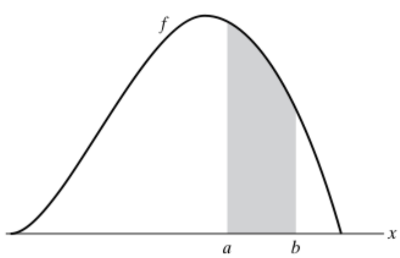
↳ Random variables whose set of possible values is either finite or countably infinite.

However, there are random variables whose set of possible values is uncountable.

→ We say that X is a continuous [absolutely continuous] random variable, if there exists a non-negative function f, defined for all real $x \in (-\infty, \infty)$, for any set B of real numbers.

$$\boxed{P\{X \in B\} = \int_B f(x)\,dx} \quad (1.1)$$

probability density function of a random variable X.



$P(a \le X \le b)$ = area of shaded region

FIGURE 5.1: Probability density function f.

→ In words (1.1) states that the probability that X will be in B may be obtained by integrating the probability density function over the set B.

Since X must assume some value

f must satisfy.

$$\boxed{1 = P\{X \in (-\infty, \infty)\} = \int_{-\infty}^{\infty} f(x)\,dx.}$$

- Taking Figure 5.1.

$$P\{a \le X \le b\} = \int_a^b f(x)\,dx.$$

If we let $a=b$

$$\boxed{P(X=a) = \int_a^a f(x)\,dx = 0 \quad (1.2)}$$

In other words, this equation (1.2), states that the probability that a continuous random variable will assume any fixed value is zero

↳ Hence, for a continuous random variable,

$$\boxed{P\{X < a\} = P\{X \le a\} = F(a) = \int_{-\infty}^{a} f(x)\,dx}$$

EXAMPLE 1

Suppose X is a continuous random variable whose probability density function is given by

$$f(x) = \begin{cases} C(4x - 2x^2) & 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

D What is the value of C?

$$c\int_0^2 (4x-2x^2)\,dx = 1$$

$$c\left[2x^2-\frac{2x^3}{3}\right]_0^2 = 1$$

$$c\left[\left[2(2)^2-\frac{2(2)^3}{3}\right]-\left[0-0\right]\right]=1$$

$$c\left[\frac{8}{3}\right]=1 \iff C=3/8 \;//$$

b) $P(X>1)=\int_1^{\infty} f(x)\,dx$

$$=\frac{3}{8}\int_1^2 (4x-2x^2)\,dx = 1/2 \;//$$

5.2 - EXPECTATION and Variance of continuous random variables

Reminder:

In Chapter 4, we defined the expected value of a discrete random variable X, by

$$E[X] = \sum_x x P(X = x)$$

If X is a continuous random variable having probability density function $f(x)$ then, because

$f(x)dx \approx P(x \leq X \leq x + dx)$ for (dx small)

it is easy to see that the analogous definition is to define the expected value of X by

$$E[X] = \int_{-\infty}^{\infty} x f(x)\, dx.$$

Example 2a

Find $E[X]$ when the density function of X is

$$f(x) = \begin{cases} 2x & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

$\hookrightarrow$ $E[X] = \int x f(x) dx$

$$= \int_0^1 2x^2\, dx = 2/3$$

Example 2b

The density function of x is given by

$$f(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find $E[e^x]$

→ Let $Y = e^x$

↳ S1 Probability distribution function of Y

$$\rightarrow F_Y(x) = P(Y \leq x)$$
$$= P(e^x \leq x)$$
$$= P(X \leq \log(x))$$
$$= \int_0^{\log(x)} f(y)\,dy$$
$$= \log(x)$$

By differentiating $F_Y(x)$, we can conclude that the probability density function of Y is given by

$$f_Y(x) = \frac{1}{x} \qquad 1 \le x \le e$$

Hence

$$E[e^Y] = E[Y] = \int_{-\infty}^{\infty} x \, f_Y(x)\, dx$$

$$= \int_1^e dx$$

$$= \underline{\underline{e - 1}}$$

Although the method employed in Example 2b to compute the expected value of a function of X is always applicable, there is, as in the discrete case, an alternative way of proceeding. The following is a direct analog of Proposition 4.1. of Chapter 4.

Proposition 2.1 - Important

If X is a continuous random variable with probability density function $f(x)$, then, for any real-valued function g

$$\boxed{E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\,dx.}$$

↳ An application of Proposition 2.1 to (example 2b) yields

$$E[e^X] = \int_0^1 e^x \quad \text{since } f(x) = 1 \quad 0 < x < 1$$

$$= e - 1$$

Proof

Lemma 2.1
For a nonnegative random variable Y,

$$E[Y] = \int_0^\infty P\{Y > y\}\, dy$$

Continuous Random Variables

Proof. We present a proof when Y is a continuous random variable with probability density function f_Y. We have

$$\int_0^\infty P\{Y > y\}\, dy = \int_0^\infty \int_y^\infty f_Y(x)\, dx\, dy$$

where we have used the fact that $P\{Y > y\} = \int_y^\infty f_Y(x)\, dx$. Interchanging the order of integration in the preceding equation yields

$$\begin{aligned}\int_0^\infty P\{Y > y\}\, dy &= \int_0^\infty \left(\int_0^x dy\right) f_Y(x)\, dx \\ &= \int_0^\infty x f_Y(x)\, dx \\ &= E[Y] \qquad \blacksquare\end{aligned}$$

Proof of Proposition 2.1. From Lemma 2.1, for any function g for which $g(x) \geq 0$,

$$\begin{aligned}E[g(X)] &= \int_0^\infty P\{g(X) > y\}\, dy \\ &= \int_0^\infty \int_{x:g(x)>y} f(x)\, dx\, dy \\ &= \int_{x:g(x)>0} \int_0^{g(x)} dy\, f(x)\, dx \\ &= \int_{x:g(x)>0} g(x) f(x)\, dx\end{aligned}$$

which completes the proof.

EXAMPLE 2e

Find Var(X) for X as given in Example 2a.

Solution. We first compute $E[X^2]$.

$$
\begin{aligned}
E[X^2] &= \int_{-\infty}^{\infty} x^2 f(x)\,dx \\
&= \int_0^1 2x^3\,dx \\
&= \frac{1}{2}
\end{aligned}
$$

Hence, since $E[X] = \frac{2}{3}$, we obtain

$$
\text{Var}(X) = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18} \qquad \blacksquare
$$

It can be shown that, for constants a and b,

$$
\text{Var}(aX + b) = a^2\text{Var}(X)
$$

The proof mimics the one given for discrete random variables.

There are several important classes of continuous random variables that appear frequently in applications of probability; the next few sections are devoted to a study of some of them.

Corollary 2.1. If a and b are constants, then

$$
E[aX + b] = aE[X] + b
$$

Example 2d

Continuous Random Variables

The proof of Corollary 2.1 for a continuous random variable X is the same as the one given for a discrete random variable. The only modification is that the sum is replaced by an integral and the probability mass function by a probability density function.

The variance of a continuous random variable is defined exactly as it is for a discrete random variable, namely, if X is a random variable with expected value μ, then the variance of X is defined (for any type of random variable) by

$$
\text{Var}(X) = E[(X - \mu)^2]
$$

The alternative formula,

$$
\text{Var}(X) = E[X^2] - (E[X])^2
$$

is established in a manner similar to its counterpart in the discrete case.

5.3 - UNIFORM RANDOM VARIABLE

A random variable is said to be uniformly distributed over the interval $(0, 1)$ if its probability density function is given by:

$$\boxed{f(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases} \qquad 3.1}$$

Note:

Equation (3.1) is a density function since ① $f(x) \geq 0$

② $\int_{-\infty}^{\infty} f(x)\,dx = \int_0^1 dx = 1$

→ Because $f(x) > 0$ only when $x \in (0, 1)$. It follows that X must assume a value in interval $(0,1)$

$\Rightarrow$ Also $f(x)$ is constant for $x \in (0,1)$, X is just as likely to be near any value in $(0,1)$ as it is to be near any other value.

$\hookrightarrow$ Verify this statement., note that for any $0 < a < b < 1$

$$P\{a \leq X \leq b\} = \int_a^b f(x)\,dx = b - a$$

In other words, the probability that X is in any particular subinterval of $(0,1)$ equals the length of that subinterval

In general we say that X is a uniform random variable on the interval (α, β) if the probability density function of X is given by

$$f(x) = \begin{cases} \frac{1}{\beta - \alpha} & \text{if } a < x < b \\ 0 & \text{otherwise} \end{cases} \qquad (3.2)$$

Since $F(a) = \int_{-\infty}^{a} f(x)\,dx$, it follows from (3.2), that the distribution function of a uniform random variable on the interval (α, β) is given by:

$$0 \qquad a \leq \alpha$$

$$f(x) = \begin{cases} \dfrac{a-\alpha}{\beta-\alpha} & \alpha < a < \beta \\ 1 & a \geq \beta \end{cases}$$

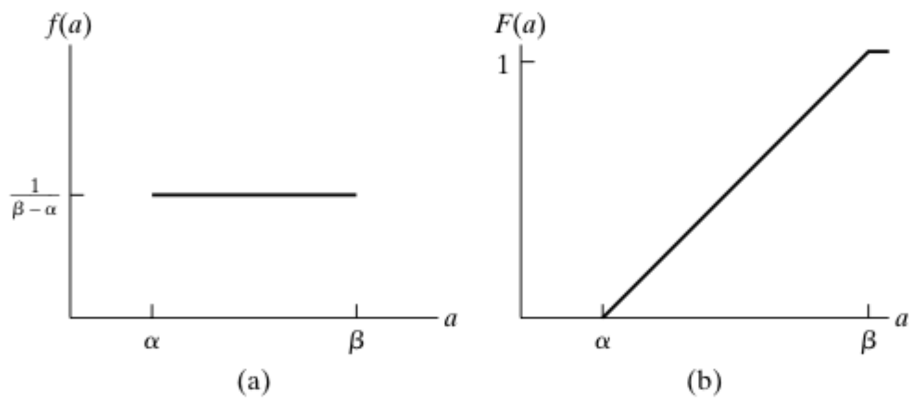


FIGURE 5.3: Graph of (a) $f(a)$ and (b) $F(a)$ for a uniform (α, β) random variable.

EXAMPLE 3a

Let X be uniformly distributed over (α, β)

- Find $E[X]$

$$E[X] = \int_{-\infty}^{\infty} x f(x)\, dx$$

$$= \int^{\beta} x \left[\frac{1}{\quad} \right] dx$$

$$= \frac{1}{b-a}\int_a^b x\,dx = \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b$$

$$= \frac{1}{2(b-a)}\left(b^2 - a^2\right)$$

$$= \frac{\cancel{(b-a)}(b+a)}{2\cancel{(b-a)}} = \frac{b+a}{2} \quad //$$

→ In other words, the expected value of a random variable that is uniformly distributed over some interval is equal to the midpoint of the interval.

- $Var(X)$

↳ Need to first calculate $E[X^2]$

$$= E[x^2] = \int_a^b \frac{1}{b-a} x^2 \, dx$$

$$= \frac{1}{b-a} \int_a^b x^2 \, dx$$

$$= \frac{1}{3(b-a)} \left[\frac{x^3}{3} \right]_a^b = \frac{1}{3(b-a)} (b^3 - a^3)$$

$$= \frac{b^2 + ab + a^2}{3}$$

Hence

$$Var(X) = E[X^2] - [E[X]]^2$$

$$= \frac{b^2 + ab + a^2}{3} - \frac{(a+b)^2}{4}$$

$$= \frac{(b-a)^2}{12} \quad //$$

Therefore, the variance of a random variable that is uniformly distributed over some interval is the square of the length of that interval divided by 12.

Example 3b

- X is uniformly distributed over $(0, 10)$ calculate the probability that

a) $X < 3$

$$P(X<3) = \int_0^3 \frac{1}{10}\,dx = \frac{3}{10}$$

b) $$P(X>6) = \int_6^{10} \frac{1}{10}\,dx = \frac{4}{10}$$

(c) $$P(3<X<8) = \int_3^8 \frac{1}{10}\,dx = \frac{1}{2}$$

Alice example.

The next example was first considered by the French mathematician Joseph L. F. Bertrand in 1889 and is often referred to as *Bertrand's paradox*. It represents our initial introduction to a subject commonly referred to as geometrical probability.

EXAMPLE 3d

Consider a random chord of a circle. What is the probability that the length of the chord will be greater than the side of the equilateral triangle inscribed in that circle?

Solution. As stated, the problem is incapable of solution because it is not clear what is meant by a random chord. To give meaning to this phrase, we shall reformulate the problem in two distinct ways.

The first formulation is as follows: The position of the chord can be determined by its distance from the center of the circle. This distance can vary between 0 and r, the radius of the circle. Now, the length of the chord will be greater than the side of the equilateral triangle inscribed in the circle if the distance from the chord to the center of the circle is less than $r/2$. Hence, by assuming that a random chord is a chord whose distance D from the center of the circle is uniformly distributed between 0 and r, we see that the probability that the length of the chord is greater than the side of an inscribed equilateral triangle is

$$P\left\{D < \frac{r}{2}\right\} = \frac{r/2}{r} = \frac{1}{2}$$

For our second formulation of the problem, consider an arbitrary chord of the circle; through one end of the chord, draw a tangent. The angle θ between the chord and the tangent, which can vary from $0°$ to $180°$, determines the position of the chord. (See Figure 5.4.) Furthermore, the length of the chord will be greater than the side of the inscribed equilateral triangle if the angle θ is between $60°$ and $120°$. Hence, assuming that a random chord is a chord whose angle θ is uniformly distributed between $0°$ and $180°$, we see that the desired answer in this formulation is

$$P\{60 < \theta < 120\} = \frac{120 - 60}{180} = \frac{1}{3}$$

Note that random experiments could be performed in such a way that $\frac{1}{2}$ or $\frac{1}{3}$ would be the correct probability. For instance, if a circular disk of radius r is thrown on a table ruled with parallel lines a distance $2r$ apart, then one and only one of these lines would cross the disk and form a chord. All distances from this chord to the center of the disk would be equally likely, so that the desired probability that the chord's length will be greater than the side of an inscribed equilateral triangle is $\frac{1}{2}$. In contrast, if the

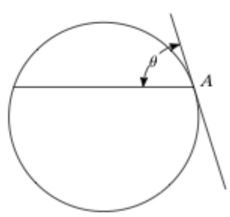


FIGURE 5.4

experiment consisted of rotating a needle freely about a point A on the edge (see Figure 5.4) of the circle, the desired answer would be $\frac{1}{3}$. ■

5.4 - NORMAL RANDOM VARIABLES

We say that X is a normal random variable / normally distributed, with parameters μ & σ^2 if the density of X is given by:

$$\boxed{f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(x-\mu)^2/2\sigma^2} \qquad -\infty < x < \infty}$$

This density function is a bell-shaped curve that is symmetric about μ. (See Figure 5.5.)

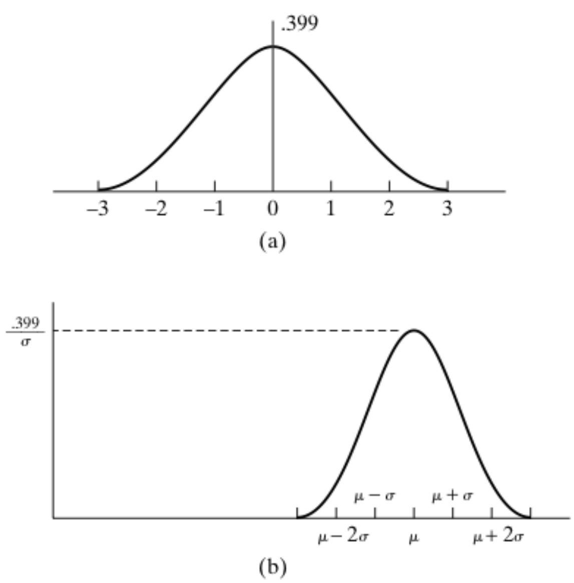


FIGURE 5.5: Normal density function: (a) $\mu = 0, \sigma = 1$; (b) arbitrary μ, σ^2.

To prove that $f(x)$ is indeed a probability density function, we need to show that

$$\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2}\,dx = 1$$

†The other is the strong law of large numbers.

Making the substitution $y = (x - \mu)/\sigma$, we see that

$$\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2}\,dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-y^2/2}\,dy$$

Hence, we must show that

$$\int_{-\infty}^{\infty} e^{-y^2/2}\,dy = \sqrt{2\pi}$$

Toward this end, let $I = \int_{-\infty}^{\infty} e^{-y^2/2}\,dy$. Then

$$\begin{aligned}
I^2 &= \int_{-\infty}^{\infty} e^{-y^2/2}\,dy \int_{-\infty}^{\infty} e^{-x^2/2}\,dx \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(y^2+x^2)/2}\,dy\,dx
\end{aligned}$$

We now evaluate the double integral by means of a change of variables to polar coordinates. (That is, let $x = r\cos\theta$, $y = r\sin\theta$, and $dy\,dx = r\,d\theta\,dr$.) Thus,

$$\begin{aligned}
I^2 &= \int_0^{\infty}\int_0^{2\pi} e^{-r^2/2} r\,d\theta\,dr \\
&= 2\pi\int_0^{\infty} re^{-r^2/2}dr \\
&= -2\pi e^{-r^2/2}\Big|_0^{\infty} \\
&= 2\pi
\end{aligned}$$

Hence, $I = \sqrt{2\pi}$, and the result is proved.

An important fact about normal random variables is that if X is normally distributed with parameters μ and σ^2, then $Y = aX + b$ is normally distributed with parameters $a\mu + b$ and $a^2\sigma^2$. To prove this statement, suppose that $a > 0$. (The proof when $a < 0$ is similar.) Let F_Y denote the cumulative distribution function of Y. Then

$$\begin{aligned}
F_Y(x) &= P\{Y \le x\} \\
&= P\{aX + b \le x\} \\
&= P\{X \le \frac{x - b}{a}\} \\
&= F_X(\frac{x - b}{a})
\end{aligned}$$

where F_X is the cumulative distribution function of X. By differentiation, the density function of Y is then

$$
\begin{aligned}
f_Y(x) &= \frac{1}{a} f_X\left(\frac{x-b}{a}\right) \\
&= \frac{1}{\sqrt{2\pi}\, a\sigma} \exp\{-\left(\frac{x-b}{a} - \mu\right)^2 / 2\sigma^2\} \\
&= \frac{1}{\sqrt{2\pi}\, a\sigma} \exp\{-(x - b - a\mu)^2 / 2(a\sigma)^2\}
\end{aligned}
$$

which shows that Y is normal with parameters $a\mu + b$ and $a^2\sigma^2$.

An important implication of the preceding result is that if X is normally distributed with parameters μ and σ^2, then $Z = (X - \mu)/\sigma$ is normally distributed with parameters 0 and 1. Such a random variable is said to be a *standard*, or a *unit*, normal random variable.

We now show that the parameters μ and σ^2 of a normal random variable represent, respectively, its expected value and variance.

(

EXAMPLE 4a

Find $E[X]$ and $Var(X)$ when X is a normal random variable with parameters μ and σ^2

Solution

Lets start by finding the mean & variance of the standard normal random variable

$Z = (X-\mu)/\sigma$

$$E[Z] = \int_{-\infty}^{\infty} x f_Z(x)\, dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-x^2/2}\, dx$$

$$= -\frac{1}{\sqrt{2\pi}} e^{-x^2/2} \Big|_{-\infty}^{\infty}$$

$$= 0 \;//$$

Thus

$$Var(Z) = E[Z^2]$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-x^2/2} \, dx$$

- Integration by parts (with $u = x$ & $dv = xe^{-x^2/2}$)

↳ $\underline{\underline{Var(Z)}}$

$$Var(Z) = \frac{1}{\sqrt{2\pi}} \left(-xe^{-x^2/2} \Big|_{-\infty}^{\infty} \right) + \int_{-\infty}^{\infty} e^{-x^2/2} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} dx$$

$$= 1$$

Because $X = \mu + \sigma Z$, the preceeding yields the results:

- $E[X] = \mu + \sigma E[Z] = \mu$
- $Var(X) = \sigma^2 Var(Z) = \sigma^2$

- Cumulative distribution function of a standard normal random variable by $\Phi(x)$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2}\, dy$$

- TABLE of NORMAL.

TABLE 5.1: AREA $\Phi(x)$ UNDER THE STANDARD NORMAL CURVE TO THE LEFT OF X

X	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

- The values of $\Phi(x)$ for <u>non-negative</u> x

- For negative values of x, $\phi(x)$ can be obtained from relationship

$$\Phi(-x) = 1 - \Phi(x) \quad -\infty < x < \infty$$

This equation states that if Z is a standard normal variable, then:

$$P(Z \leq -x) = P(Z > x) \quad -\infty < x < \infty$$

Since $Z = \frac{(X - \mu)}{\sigma}$ is a standard normal random variable, whenever

X is normally distributed with parameters μ and σ^2, it follows that the distribution function of X can be expressed a:

$$F_X(a) = P(X \leq a) = P\left(\frac{X-\mu}{\sigma} \leq \frac{a-\mu}{\sigma}\right)$$

$$= \Phi\left(\frac{a-\mu}{\sigma}\right)$$

5.4.1 - The Normal Approximation to the Binomial Distribution

* DeMoivre-Laplace limit theorem states that when n is large, a binomial random variable with parameters n & p will have

"approximately" the same distribution as a normal random variable with the same mean & variance as the binomial

The DeMoivre–Laplace limit theorem

If S_n denotes the number of successes that occur when n independent trials, each resulting in a success with probability p, are performed, then, for any $a < b$,

$$P\left\{a \leq \frac{S_n - np}{\sqrt{np(1 - p)}} \leq b\right\} \rightarrow \Phi(b) - \Phi(a)$$

as $n \rightarrow \infty$.

* **Note:** *

Now we have 2 possible approximation to binomial probabilities:

① Poisson approximation.

↳ good when $\boxed{\begin{array}{l} n \text{ is large} \\ p \text{ is small} \end{array}}$

② Normal approximation

↳ good when {$np(1-p)$ is large}

$$\{np(1-p) \geq 10\}$$

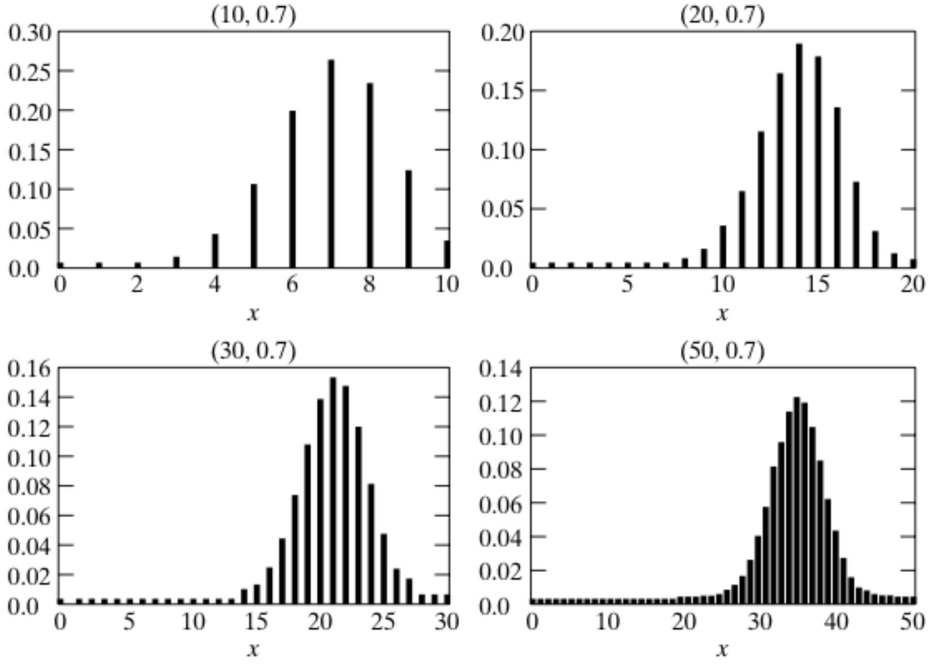


FIGURE 5.6: The probability mass function of a binomial (n, p) random variable becomes more and more "normal" as n becomes larger and larger.

EXAMPLE 4h

To determine the effectiveness of a certain diet in reducing the amount of cholesterol in the bloodstream, 100 people are put on the diet. After they have been on the diet for a sufficient length of time, their cholesterol count will be taken. The nutritionist running this experiment has decided to endorse the diet if at least 65 percent of the people have a lower cholesterol count after going on the diet. What is the probability that the nutritionist endorses the new diet if, in fact, it has no effect on the cholesterol level?

Solution. Let us assume that if the diet has no effect on the cholesterol count, then, strictly by chance, each person's count will be lower than it was before the diet with probability $\frac{1}{2}$. Hence, if X is the number of people whose count is lowered, then the probability that the nutritionist will endorse the diet when it actually has no effect on the cholesterol count is

$$\begin{aligned}
\sum_{i=65}^{100} \binom{100}{i} \left(\frac{1}{2}\right)^{100} &= P\{X \geq 64.5\} \\
&= P\left\{\frac{X - (100)(\frac{1}{2})}{\sqrt{100(\frac{1}{2})(\frac{1}{2})}} \geq 2.9\right\} \\
&\approx 1 - \Phi(2.9) \\
&\approx .0019
\end{aligned}$$

■

5.5 - Exponential Random Variable

A continuous random variable whose probability density function is given for some $\lambda > 0$, by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

is said to be exponential random variable with parameter λ.

The cumulative distribution function $F(a)$ of an exponential random variable is given by

$$F(a) = P(X \leq a)$$
$$= \int_0^a \lambda e^{-\lambda x} dx$$
$$= -e^{-\lambda x}\Big|_0^a$$
$$= 1 - e^{-\lambda a} \quad a \geq 0$$

Note

That $F(\infty) = \int_0^\infty \lambda e^{-\lambda x} dx = 1$

$\therefore$ now the parameter λ will now be shown to equal the reciprocal of the expected value.

* Example 5a

Let X be an exponential random variable with parameters λ.

- Calculate $E[X]$.

Since the density function is given by:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

we obtain, for $n > 0$

$$E[X^n] = \int_0^\infty x^n \lambda e^{-\lambda x} dx$$

Integrating by parts. $\left(\begin{matrix} \lambda e^{-\lambda x} = dv \\ u = x^n \end{matrix} \right)$

$$E[X^n] = -x^n e^{-\lambda x} \Big|_0^\infty + \int_0^\infty e^{-\lambda x} n x^{n-1} dx$$

$$= 0 + \frac{n}{\lambda}\int_0^{\infty} \lambda e^{-\lambda x} x^{n-1}\, dx$$

$$\boxed{= \frac{n}{\lambda} E[X^{n-1}]}$$

Letting $n=1$ & then $n=2$ gives:

$$\boxed{E[X] = \frac{1}{\lambda}}$$

$$E[X^2] = \frac{2}{\lambda} E[X] = \frac{2}{\lambda^2}$$

$\therefore$ Variance:

$$\boxed{\mathrm{Var}(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}}$$

$\therefore$ Thus, the mean of the exponential is the reciprocal of its parameter λ, & the variance is the mean squared.

In practice, the exponential distribution often arises as the distribution of the amount of time until some specific event occurs

- Amount of time (starting now) until an earthquake occurs.

EXAMPLE 5b

Suppose that the length of a phone call in minutes is an exponential random variable with parameter $\lambda = \frac{1}{10}$. If someone arrives immediately ahead of you at a public telephone booth, find the probability that you will have to wait

(a) more than 10 minutes;
(b) between 10 and 20 minutes.

Solution. Let X denote the length of the call made by the person in the booth. Then the desired probabilities are

(a)

$$\begin{aligned} P\{X > 10\} &= 1 - F(10) \\ &= e^{-1} \approx .368 \end{aligned}$$

(b)

$$\begin{aligned} P\{10 < X < 20\} &= F(20) - F(10) \\ &= e^{-1} - e^{-2} \approx .233 \end{aligned}$$ ■

We say that a nonnegative random variable X is *memoryless* if

$$P\{X > s + t \,|\, X > t\} = P\{X > s\} \quad \text{for all } s, t \geq 0 \tag{5.1}$$

If we think of X as being the lifetime of some instrument, Equation (5.1) states that the probability that the instrument survives for at least $s+t$ hours, given that it has survived t hours, is the same as the initial probability that it survives for at least s hours. In other words, if the instrument is alive at age t, the distribution of the remaining amount of time that it survives is the same as the original lifetime distribution. (That is, it is as if the instrument does not "remember" that it has already been in use for a time t.)

Equation (5.1) is equivalent to

$$\frac{P\{X > s + t, X > t\}}{P\{X > t\}} = P\{X > s\}$$

or

$$P\{X > s + t\} = P\{X > s\}P\{X > t\} \tag{5.2}$$

Since Equation (5.2) is satisfied when X is exponentially distributed (for $e^{-\lambda(s+t)} = e^{-\lambda s}e^{-\lambda t}$), it follows that exponentially distributed random variables are memoryless.

5.5.1 - Hazard Rate Function

Consider a positive continuous random variable X that we interpret as being the lifetime of some item. Let X have distribution function F and density f.

The hazard rate / failure rate function $\lambda(t)$ of F is defined by

$$\lambda(t) = \frac{f(t)}{\bar{F}(t)}, \text{ where } \bar{F} = 1 - F$$

To interpret $\lambda(t)$, suppose that the item has survived for a time (t) and we desire the probability that it will not survive for an additional time (dt). That is, $P(X \in t, t+dt) \mid X > t)$

$$P\{X \in (t, t+dt) \mid X > t\} = \frac{P\{X \in (t, t+dt), X > t\}}{P\{X > t\}}$$
$$= \frac{P\{X \in (t, t+dt)\}}{P\{X > t\}}$$
$$\approx \frac{f(t)}{\overline{F}(t)}\, dt$$

Thus $\lambda(t)$ represents the conditional probability intensity that a t-unit-old item will fail.

- Suppose now that the lifetime distribution is exponential. Then by the memoryless property, it follows that the distribution of remaining life for a t-years-old item is the same as that for a new item.

Hence, $\lambda(t)$ should be constant

$$\lambda(t) = \frac{f(t)}{\bar{F}(t)} = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda$$

→ Thus failure rate function for the exponential distribution is constant.

∞ The parameter λ is often referred to as the rate of the distribution.

It turns out that the failure rate function $\lambda(t)$ uniquely determines the distribution F. To prove this, note that, by definition,

$$\lambda(t) = \frac{\frac{d}{dt}F(t)}{1 - F(t)}$$

Integrating both sides yields

$$\log(1 - F(t)) = -\int_0^t \lambda(t)\,dt + k$$

or

$$1 - F(t) = e^k \exp\left\{-\int_0^t \lambda(t)\,dt\right\}$$

Letting $t = 0$ shows that $k = 0$; thus,

$$F(t) = 1 - \exp\left\{-\int_0^t \lambda(t)\,dt\right\} \tag{5.4}$$

Hence, a distribution function of a positive continuous random variable can be specified by giving its hazard rate function. For instance, if a random variable has a linear hazard rate function—that is, if

$$\lambda(t) = a + bt$$

then its distribution function is given by

$$F(t) = 1 - e^{-at-bt^2/2}$$

and differentiation yields its density, namely,

$$f(t) = (a + bt)e^{-(at+bt^2/2)} \quad t \geq 0$$

When $a = 0$, the preceding equation is known as the *Rayleigh density function*.

EXAMPLE 5f

EXAMPLE 5f

One often hears that the death rate of a person who smokes is, at each age, twice that of a nonsmoker. What does this mean? Does it mean that a nonsmoker has twice the probability of surviving a given number of years as does a smoker of the same age?

Solution. If $\lambda_s(t)$ denotes the hazard rate of a smoker of age t and $\lambda_n(t)$ that of a nonsmoker of age t, then the statement at issue is equivalent to the statement that

$$\lambda_s(t) = 2\lambda_n(t)$$

The probability that an A-year-old nonsmoker will survive until age B, $A < B$, is

$$\begin{aligned}
&P\{A\text{-year-old nonsmoker reaches age } B\} \\
&\quad = P\{\text{nonsmoker's lifetime} > B \mid \text{nonsmoker's lifetime} > A\} \\
&\quad = \frac{1 - F_{\text{non}}(B)}{1 - F_{\text{non}}(A)} \\
&\quad = \frac{\exp\left\{-\int_0^B \lambda_n(t)\,dt\right\}}{\exp\left\{-\int_0^A \lambda_n(t)\,dt\right\}} \qquad \text{from (5.4)} \\
&\quad = \exp\left\{-\int_A^B \lambda_n(t)\,dt\right\}
\end{aligned}$$

whereas the corresponding probability for a smoker is, by the same reasoning,

$$\begin{aligned}
P\{A\text{-year-old smoker reaches age } B\} &= \exp\left\{-\int_A^B \lambda_s(t)\,dt\right\} \\
&= \exp\left\{-2\int_A^B \lambda_n(t)\,dt\right\} \\
&= \left[\exp\left\{-\int_A^B \lambda_n(t)\,dt\right\}\right]^2
\end{aligned}$$

In other words, for two people of the same age, one of whom is a smoker and the other a nonsmoker, the probability that the smoker survives to any given age is the *square* (not one-half) of the corresponding probability for a nonsmoker. For instance, if $\lambda_n(t) = \frac{1}{30}$, $50 \le t \le 60$, then the probability that a 50-year-old nonsmoker reaches age 60 is $e^{-1/3} \approx .7165$, whereas the corresponding probability for a smoker is $e^{-2/3} \approx .5134$. ■

Important

5.6 - OTHER CONTINUOUS DISTRIBUTION

5.6.1 - The Gamma Distribution

A random variable is said to have a gamma distribution with parameters (a, λ), $\lambda > 0$, $a > 0$ if its density function is given by:

$$f(x) = \begin{cases} \dfrac{\lambda e^{-\lambda x} (\lambda x)^{a-1}}{\Gamma a} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

where $\Gamma(a)$ is gamma function

$$\Gamma(a) = \int_0^\infty e^{-y} y^{a-1} \, dy$$

Integration of $\Gamma(\alpha)$ by parts yields

$$\begin{aligned}
\Gamma(\alpha) &= -e^{-y}y^{\alpha-1}\Big|_0^\infty + \int_0^\infty e^{-y}(\alpha - 1)y^{\alpha-2}\,dy \\
&= (\alpha - 1)\int_0^\infty e^{-y}y^{\alpha-2}\,dy \qquad (6.1)\\
&= (\alpha - 1)\Gamma(\alpha - 1)
\end{aligned}$$

For integral values of α, say, $\alpha = n$, we obtain, by applying Equation (6.1) repeatedly,

$$\begin{aligned}
\Gamma(n) &= (n - 1)\Gamma(n - 1) \\
&= (n - 1)(n - 2)\Gamma(n - 2) \\
&= \cdots \\
&= (n - 1)(n - 2)\cdots 3 \cdot 2\Gamma(1)
\end{aligned}$$

Since $\Gamma(1) = \int_0^\infty e^{-x}\,dx = 1$, it follows that, for integral values of n,

$$\Gamma(n) = (n - 1)!$$

When α is a positive integer, say, $\alpha = n$, the gamma distribution with parameters (α, λ) often arises, in practice as the distribution of the amount of time one has to wait until a total of n events has occurred. More specifically, if events are occurring randomly and in accordance with the three axioms of Section 4.7, then it turns out that the amount of time one has to wait until a total of n events has occurred will be a gamma random variable with parameters (n, λ). To prove this, let T_n denote the time at which the nth event occurs, and note that T_n is less than or equal to t if and only if the number of events that have occurred by time t is at least n. That is, with $N(t)$ equal to the number of events in $[0, t]$,

$$\begin{aligned}
P\{T_n \le t\} &= P\{N(t) \ge n\} \\
&= \sum_{j=n}^{\infty} P\{N(t) = j\} \\
&= \sum_{j=n}^{\infty} \frac{e^{-\lambda t}(\lambda t)^j}{j!}
\end{aligned}$$

where the final identity follows because the number of events in $[0, t]$ has a Poisson distribution with parameter λt. Differentiation of the preceding now yields the density function of T_n:

$$\begin{aligned}
f(t) &= \sum_{j=n}^{\infty} \frac{e^{-\lambda t} j(\lambda t)^{j-1}\lambda}{j!} - \sum_{j=n}^{\infty} \frac{\lambda e^{-\lambda t}(\lambda t)^j}{j!} \\
&= \sum_{j=n}^{\infty} \frac{\lambda e^{-\lambda t}(\lambda t)^{j-1}}{(j - 1)!} - \sum_{j=n}^{\infty} \frac{\lambda e^{-\lambda t}(\lambda t)^j}{j!} \\
&= \frac{\lambda e^{-\lambda t}(\lambda t)^{n-1}}{(n - 1)!}
\end{aligned}$$

Hence, T_n has the gamma distribution with parameters (n, λ). (This distribution is often referred to in the literature as the *n-Erlang distribution*.) Note that when $n = 1$, this distribution reduces to the exponential distribution.

The gamma distribution with $\lambda = \frac{1}{2}$ and $\alpha = n/2$, n a positive integer, is called the χ_n^2 (read "chi-squared") distribution with n degrees of freedom. The chi-squared distribution often arises in practice as the distribution of the error involved in attempting to hit a target in n-dimensional space when each coordinate error is normally distributed. This distribution will be studied in Chapter 6, where its relation to the normal distribution is detailed.

EXAMPLE 6a

Let X be a gamma random variable with parameters α and λ. Calculate (a) $E[X]$ and (b) $\text{Var}(X)$.

Solution. (a)

$$\begin{aligned}
E[X] &= \frac{1}{\Gamma(\alpha)}\int_0^\infty \lambda x e^{-\lambda x}(\lambda x)^{\alpha-1}\,dx \\
&= \frac{1}{\lambda\Gamma(\alpha)}\int_0^\infty \lambda e^{-\lambda x}(\lambda x)^{\alpha}\,dx \\
&= \frac{\Gamma(\alpha+1)}{\lambda\Gamma(\alpha)} \\
&= \frac{\alpha}{\lambda} \quad \text{by Equation (6.1)}
\end{aligned}$$

(b) By first calculating $E[X^2]$, we can show that

$$\text{Var}(X) = \frac{\alpha}{\lambda^2}$$

The details are left as an exercise. ■

5.6.2 - The Weibull Distribution

The Weibull distribution is used in engineering practice due to its versatility.

→ It is widely used in the field of life phenomena as the distribution of the lifetime of some objects, especially when the "weakest link" model is appropriate for the object.

- Consider for example an object consisting of many parts, and suppose that the object experiences death (failure) when any of its parts fail.

It has been shown (both theoretically

& empirically) that under these conditions a Weibull distribution provides a close approximation to the distribution of the lifetime of the item.

* The Weibull distribution function

$$F(x) = \begin{cases} 0 & x \le v \\ 1 - \exp\left\{ -\left(\frac{x - v}{a} \right)^{\beta} \right\} & x > v \end{cases}$$

(6.2)

- A random variable whose cumulative distribution function is given by Equation (6.2) is said to be Weibull random variable with parameters v, a, β

↳ Differentiation yields the density:

$$f(x) = \begin{cases} 0 & x \le \nu \\ \frac{\beta}{\alpha}\left(\frac{x-\nu}{\alpha}\right)^{\beta-1} \exp\left\{-\left(\frac{x-\nu}{\alpha}\right)^{\beta}\right\} & x > \nu \end{cases}$$

5.6.3 - The Cauchy Distribution

A random variable is said to have Cauchy distribution with parameter θ, $-\infty < \theta < \infty$, if its density is given by:

$$f(x) = \frac{1}{\pi}\frac{1}{1+(x-\theta)^2} \qquad -\infty < x < \infty$$

EXAMPLE 6b

Suppose that a narrow-beam flashlight is spun around its center, which is located a unit distance from the x-axis. (See Figure 5.7.) Consider the point X at which the beam intersects the x-axis when the flashlight has stopped spinning. (If the beam is not pointing toward the x-axis, repeat the experiment.)

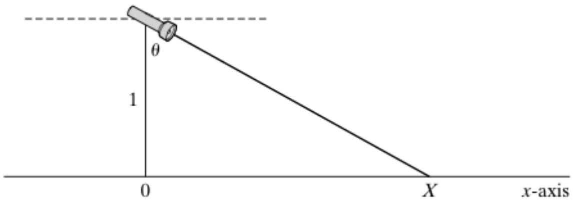


FIGURE 5.7

As indicated in Figure 5.7, the point X is determined by the angle θ between the flashlight and the y-axis, which, from the physical situation, appears to be uniformly distributed between $-\pi/2$ and $\pi/2$. The distribution function of X is thus given by

$$\begin{aligned} F(x) &= P\{X \leq x\} \\ &= P\{\tan\theta \leq x\} \\ &= P\{\theta \leq \tan^{-1} x\} \\ &= \frac{1}{2} + \frac{1}{\pi}\tan^{-1} x \end{aligned}$$

where the last equality follows since θ, being uniform over $(-\pi/2, \pi/2)$, has distribution

$$P\{\theta \leq a\} = \frac{a - (-\pi/2)}{\pi} = \frac{1}{2} + \frac{a}{\pi} \qquad -\frac{\pi}{2} < a < \frac{\pi}{2}$$

Hence, the density function of X is given by

$$f(x) = \frac{d}{dx}F(x) = \frac{1}{\pi(1 + x^2)} \qquad -\infty < x < \infty$$

and we see that X has the Cauchy distribution.† ■

5.6.4 - The Beta Distribution

A random variable is said to have a beta distribution if its density is given by

$$f(x) = \begin{cases} \dfrac{1}{B(a,b)} x^{a-1}(1-x)^{b-1} & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

where

$$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}\,dx$$

The beta distribution can be used to model a random phenomenon whose set of possible values is some finite interval $[c, d]$ - which by letting c denote the origin & taking $(d-c)$ as a unit measurement can be transformed into the interval $[0, 1]$.

- When $a = b$, the beta density is symmetric about $\frac{1}{2}$, giving more and more weight to regions about $\frac{1}{2}$ as the common value a increases

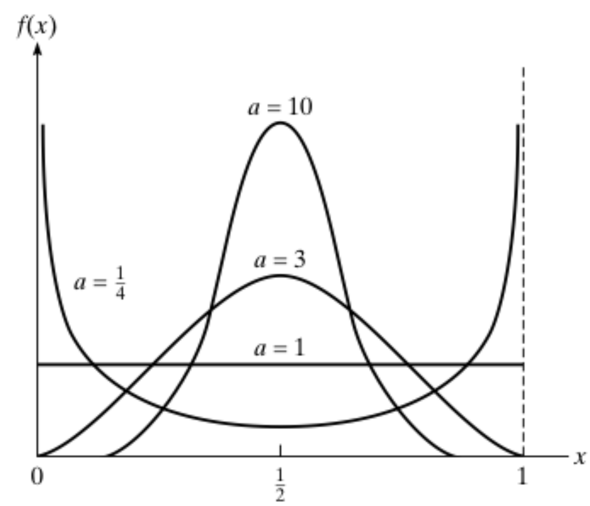


FIGURE 5.8: Beta densities with parameters (a, b) when $a = b$.

- When $b > a$, the density is skewed to the left (in the sense that smaller values become more likely).

and it is skewed to the right when $a > b$

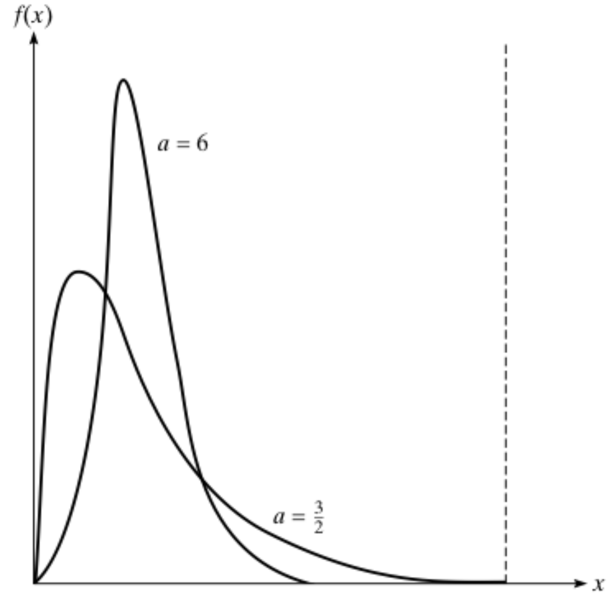


FIGURE 5.9: Beta densities with parameters (a, b) when $a/(a + b) = 1/20$.

Upon using Equation (6.1) along with the identity (6.3), it is an easy matter to show that if X is a beta random variable with parameters a and b, then

$$E[X] = \frac{a}{a + b}$$

$$\text{Var}(X) = \frac{ab}{(a + b)^2(a + b + 1)}$$

Remark. A verification of Equation (6.3) appears in Example 7c of Chapter 6. ■

Note

Note

The relationship

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \tag{6.3}$$

can be shown to exist between

$$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}\,dx$$

and the gamma function.

---

†That $\frac{d}{dx}(\tan^{-1} x) = 1/(1+x^2)$ can be seen as follows: If $y = \tan^{-1} x$, then $\tan y = x$, so

$$1 = \frac{d}{dx}(\tan y) = \frac{d}{dy}(\tan y)\frac{dy}{dx} = \frac{d}{dy}\left(\frac{\sin y}{\cos y}\right)\frac{dy}{dx} = \left(\frac{\cos^2 y + \sin^2 y}{\cos^2 y}\right)\frac{dy}{dx}$$

or

$$\frac{dy}{dx} = \frac{\cos^2 y}{\sin^2 y + \cos^2 y} = \frac{1}{\tan^2 y + 1} = \frac{1}{x^2+1}$$

5.7 - Distribution of a function of a random Variable

Often, we know the probability distribution of a random variable.

↳ and we are interested in determining the distribution of some function of it

→ To do so it is necessary to express the event that $g(x) \le y$ in terms of X being in some set

EXAMPLE 7a

Let $X \sim U(0,1)$

- $Y = X^n \therefore 0 \le y \le 1$

$$
\begin{aligned}
F_Y(y) &= P(Y \le y) \\
&= P(X^n \le y) \\
&= P(X \le y^{1/n}) \\
&= F_X(y^{1/n}) = \underline{y^{1/n}}
\end{aligned}
$$

∴ density function of Y :

$$- f_Y(y) = \begin{cases} \frac{1}{n} y^{1/n - 1} & 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

- <u>EXAMPLE 7c</u>

If X has a probability f_X

then $Y = |X|$ has a density function that is obtained as follows.

For $y \geq 0$

$$F_Y(y) = P(Y \leq y)$$

$$= P(|X| \le y)$$

$$= P(-y \le X \le y)$$

$$= F_X(y) - F_X(-y)$$

Hence, on differentiation, we obtain.

$$f_Y(y) = f_X(y) + f_X(-y) \qquad y \ge 0$$

Note

The method employed in Example 7a through 7c can be used to prove Theorem 7.1

Theorem 7.1. *Let X be a continuous random variable having probability density function f_X. Suppose that $g(x)$ is a strictly monotonic (increasing or decreasing), differentiable (and thus continuous) function of x. Then the random variable Y defined by $Y = g(X)$ has a probability density function given by*

$$f_Y(y) = \begin{cases} f_X[g^{-1}(y)] \left| \dfrac{d}{dy} g^{-1}(y) \right| & \text{if } y = g(x) \text{ for some } x \\ 0 & \text{if } y \neq g(x) \text{ for all } x \end{cases}$$

where $g^{-1}(y)$ is defined to equal that value of x such that $g(x) = y$.

We shall prove Theorem 7.1 when $g(x)$ is an increasing function.

Proof. Suppose that $y = g(x)$ for some x. Then, with $Y = g(X)$,

$$\begin{aligned} F_Y(y) &= P\{g(X) \le y\} \\ &= P\{X \le g^{-1}(y)\} \\ &= F_X(g^{-1}(y)) \end{aligned}$$

Differentiation gives

$$f_Y(y) = f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y) \qquad \blacksquare$$

which agrees with Theorem 7.1, since $g^{-1}(y)$ is nondecreasing, so its derivative is nonnegative.

When $y \neq g(x)$ for any x, then $F_Y(y)$ is either 0 or 1, and in either case $f_Y(y) = 0$.

SUMMARY

A random variable X is *continuous* if there is a nonnegative function f, called the *probability density function* of X, such that, for any set B,

$$P\{X \in B\} = \int_B f(x)\,dx$$

If X is continuous, then its distribution function F will be differentiable and

$$\frac{d}{dx}F(x) = f(x)$$

The expected value of a continuous random variable X is defined by

$$E[X] = \int_{-\infty}^{\infty} xf(x)\,dx$$

A useful identity is that, for any function g,

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)\,dx$$

As in the case of a discrete random variable, the variance of X is defined by

$$\text{Var}(X) = E[(X - E[X])^2]$$

A random variable X is said to be *uniform* over the interval (a, b) if its probability density function is given by

$$f(x) = \begin{cases} \dfrac{1}{b-a} & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

Its expected value and variance are

$$E[X] = \frac{a+b}{2} \quad \text{Var}(X) = \frac{(b-a)^2}{12}$$

A random variable X is said to be *normal* with parameters μ and σ^2 if its probability density function is given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-(x-\mu)^2/2\sigma^2} \qquad -\infty < x < \infty$$

It can be shown that

$$\mu = E[X] \quad \sigma^2 = \text{Var}(X)$$

If X is normal with mean μ and variance σ^2, then Z, defined by

$$Z = \frac{X - \mu}{\sigma}$$

is normal with mean 0 and variance 1. Such a random variable is said to be a *standard* normal random variable. Probabilities about X can be expressed in terms of probabilities about the standard normal variable Z, whose probability distribution function can be obtained either from Table 5.1 or from a website.

When n is large, the probability distribution function of a binomial random variable with parameters n and p can be approximated by that of a normal random variable having mean np and variance $np(1 - p)$.

A random variable whose probability density function is of the form

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

is said to be an *exponential* random variable with parameter λ. Its expected value and variance are, respectively,

$$E[X] = \frac{1}{\lambda} \quad \text{Var}(X) = \frac{1}{\lambda^2}$$

A key property possessed only by exponential random variables is that they are *memoryless*, in the sense that, for positive s and t,

$$P\{X > s + t | X > t\} = P\{X > s\}$$

If X represents the life of an item, then the memoryless property states that, for any t, the remaining life of a t-year-old item has the same probability distribution as the life of a new item. Thus, one need not remember the age of an item to know its distribution of remaining life.

Let X be a nonnegative continuous random variable with distribution function F and density function f. The function

$$\lambda(t) = \frac{f(t)}{1 - F(t)} \quad t \geq 0$$

is called the *hazard rate*, or *failure rate*, function of F. If we interpret X as being the life of an item, then, for small values of dt, $\lambda(t)\,dt$ is approximately the probability that a t-unit-old item will fail within an additional time dt. If F is the exponential distribution with parameter λ, then

$$\lambda(t) = \lambda \quad t \geq 0$$

In addition, the exponential is the unique distribution having a constant failure rate.

A random variable is said to have a *gamma* distribution with parameters α and λ if its probability density function is equal to

$$f(x) = \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} \quad x \geq 0$$

and is 0 otherwise. The quantity $\Gamma(\alpha)$ is called the gamma function and is defined by

$$\Gamma(\alpha) = \int_0^\infty e^{-x} x^{\alpha-1}\,dx$$

The expected value and variance of a gamma random variable are, respectively,

$$E[X] = \frac{\alpha}{\lambda} \quad \text{Var}(X) = \frac{\alpha}{\lambda^2}$$

A random variable is said to have a *beta* distribution with parameters (a, b) if its probability density function is equal to

$$f(x) = \frac{1}{B(a,b)} x^{a-1}(1 - x)^{b-1} \quad 0 \le x \le 1$$

and is equal to 0 otherwise. The constant $B(a, b)$ is given by

$$B(a,b) = \int_0^1 x^{a-1}(1 - x)^{b-1}\,dx$$

The mean and variance of such a random variable are, respectively,

$$E[X] = \frac{a}{a + b} \quad \mathrm{Var}(X) = \frac{ab}{(a + b)^2(a + b + 1)}$$